Created by Neil Yetz & Gemma Wallace

Module 10 Lab Notebook.Rmd*

```
---
title: "R Notebook for PSY 652 Module 10 Lab: Hierarchical Regression Practice"
subtitle: Gemma Wallace & Neil Yetz
output:
  html_document:
    df_print: paged
    toc: yes
  html_notebook:
    toc: yes
---


# Load libraries
```{r, message=FALSE}
library(tidyverse)
library(psych)
library(olsrr)
```

Import data
```{r,message=FALSE}
obs <- read_csv("bac_obs.csv")
```

Describe the data
```{r,message=FALSE}
describe(obs)
```
```

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	median <dbl>	trimmed <dbl>	mad <dbl>	min <dbl>	max <dbl>
id	1	200	100.50	57.88	100.50	100.50	74.13	1.00	200.00
weight	2	200	68.46	9.92	68.90	68.52	9.56	37.60	91.40
typ_drks	3	200	46.86	14.67	46.50	47.11	14.08	4.00	79.00
alcexp	4	200	4.09	0.78	4.13	4.11	0.77	2.01	6.06
pmood	5	200	5.12	1.40	5.00	5.09	1.48	1.00	9.00
absorb	6	200	4.69	0.91	4.64	4.68	0.95	2.67	6.80
alc_gm	7	200	32.79	7.73	33.00	32.76	8.90	8.00	58.00
bac	8	200	0.08	0.02	0.08	0.08	0.02	0.02	0.15

8 rows | 1-10 of 13 columns

```
# Mutate bac variable
```{r}
obs <- mutate(obs, bac100 = bac*100)
```


Hierarchical Regression Practice

Build Reduced Model A
```{r}
mod_reda <- lm(data = obs, bac100 ~ typ_drks + alc_gm)
ols_regress(mod_reda)
```
```

```
                        Model Summary
------------------------------------------------------------
R                        0.869       RMSE                1.077
R-Squared                0.754       Coef. Var          13.001
Adj. R-Squared           0.752       MSE                 1.160
Pred R-Squared           0.748       MAE                 0.791
------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                               ANOVA
--------------------------------------------------------------------
               Sum of
              Squares        DF    Mean Square       F         Sig.
--------------------------------------------------------------------
Regression    702.202         2        351.101    302.586    0.0000
Residual      228.586       197          1.160
Total         930.788       199
--------------------------------------------------------------------

                                Parameter Estimates
-------------------------------------------------------------------------------------
      model     Beta    Std. Error    Std. Beta      t        Sig      lower    upper
-------------------------------------------------------------------------------------
(Intercept)    0.278         0.335                 0.830    0.408    -0.382    0.937
   typ_drks    0.012         0.007        0.078    1.634    0.104    -0.002    0.025
     alc_gm    0.228         0.013        0.814   16.962    0.000     0.201    0.254
-------------------------------------------------------------------------------------
```

```
## Build Reduced Model B
```{r}
mod_redb <- lm(data = obs, bac100 ~ typ_drks + alc_gm + weight)
ols_regress(mod_redb)
```
```

```
                        Model Summary
---------------------------------------------------------------
R                       0.961       RMSE                0.604
R-Squared               0.923       Coef. Var           7.290
Adj. R-Squared          0.922       MSE                 0.365
Pred R-Squared          0.919       MAE                 0.447
---------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                              ANOVA
----------------------------------------------------------------------
               Sum of
              Squares        DF    Mean Square       F         Sig.
----------------------------------------------------------------------
Regression    859.279         3        286.426    785.073    0.0000
Residual       71.509       196          0.365
Total         930.788       199
----------------------------------------------------------------------

                                 Parameter Estimates
-------------------------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
-------------------------------------------------------------------------------------------
(Intercept)     6.056         0.336                  18.036    0.000     5.394     6.718
   typ_drks     0.003         0.004        0.022      0.829    0.408    -0.005     0.011
     alc_gm     0.253         0.008        0.903     33.142    0.000     0.238     0.268
     weight    -0.091         0.004       -0.416    -20.749    0.000    -0.099    -0.082
-------------------------------------------------------------------------------------------
```

```
## Build Full Model
```{r}
mod_full<- lm(data = obs, bac100 ~ typ_drks + alc_gm + weight + alcexp)
ols_regress(mod_full)
```
```

```
                        Model Summary
---------------------------------------------------------------
R                       0.961       RMSE                0.605
R-Squared               0.923       Coef. Var           7.307
Adj. R-Squared          0.922       MSE                 0.367
Pred R-Squared          0.919       MAE                 0.447
---------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                              ANOVA
----------------------------------------------------------------------
               Sum of
              Squares        DF    Mean Square       F         Sig.
----------------------------------------------------------------------
Regression    859.320         4        214.830    586.161    0.0000
Residual       71.468       195          0.367
Total         930.788       199
----------------------------------------------------------------------

                                 Parameter Estimates
-------------------------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
-------------------------------------------------------------------------------------------
(Intercept)     5.984         0.400                  14.975    0.000     5.196     6.772
   typ_drks     0.003         0.004        0.019      0.683    0.495    -0.005     0.011
     alc_gm     0.252         0.008        0.900     31.164    0.000     0.236     0.268
     weight    -0.090         0.004       -0.415    -20.133    0.000    -0.099    -0.082
     alcexp     0.024         0.072        0.009      0.333    0.740    -0.119     0.167
-------------------------------------------------------------------------------------------
```

In the white space below, calculate and interpret the unique variance in Y explained by each added predictor. (Hint: think of semi-partial correlation for your interpretations)

Moving from Reduced Model A to Reduced model B increased the R^2 value from .754 to .923. This means that weight adds 16.9% in explained variance in the outcome of BAC. Moving from Model B to Model C resulted in no change in the R^2. This means that alcexp explains very little variance in BAC.

```
## Compare Reduced and Full Models

### Statistically compare model fit
```{r}
anova(mod_reda, mod_redb, mod_full, test = "F")
# list the most reduced model first, followed by the second-most reduced model, followed by the full model
```

Analysis of Variance Table

Model 1: bac100 ~ typ_drks + alc_gm
Model 2: bac100 ~ typ_drks + alc_gm + weight
Model 3: bac100 ~ typ_drks + alc_gm + weight + alcexp
  Res.Df   RSS Df Sum of Sq   F Pr(>F)
1   197 228.586
2   196 71.509 1   157.077 428.5826 <2e-16 ***
3   195 71.468 1   0.041   0.1106 0.7398

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

i. ***In the white space below, interpret the partial F-test output and answer the following questions.***

   1. ***Does adding the weight variable result in a model that explains significantly more variance in bac100 than Reduced Model A (compare Reduced Model A and Reduced Model B)?***
      - ***a.*** Yes, Model B explains a statistically significant amount more variance in BAC than model B.

   2. ***Does adding the alc_exp variable result in a model that explains significantly more variance in bac100 than Reduced Model B (compare Reduced Model B & Full Model)?***
      - a. No, Model C does NOT explain any more variance in BAC than Model C.

   3. ***Reflect on your own research interests and write 2-3 sentences describing an example of when using hierarchical regression could be a good fit for your research. When might you actually use this?***
      - a. Thanks for your answers. We enjoyed reading them 😊 – Neil & Gemma.